Acoustical analysis of Canadian French word-final vowels in varying phonetic contexts

Franzo Law II[a)]
Departments of Psychology and Communication Sciences and Disorders, University of Madison-Wisconsin, Waisman Center, 1500 Highland Avenue, Room 489, Madison, Wisconsin 53705-2280, USA
flaw@wisc.edu

Winifred Strange
Program in Speech-Language-Hearing Sciences, Graduate School and University Center-CUNY, 365 Fifth Avenue, New York, New York 10016-4309, USA
strange.win@aol.com

Abstract: This study analyzed Canadian French (CF) vowels /i y ø e ɛ o u a/ in word-final position. Of particular interest was the stability of /e-ɛ/; although some dialects of French have merged /e-ɛ/ to /e/ in word-final context, this contrast is maintained in CF. The present study investigated the stability of this contrast in various preceding phonetic contexts and in lexical vs morphological contrasts. Results showed that the contrast was maintained by all four speakers, although to varying degrees.

[AL]
Date Received: January 11, 2014 **Date Accepted:** June 8, 2015

1. Introduction

Dialects can differ greatly in their vowel production. These differences can arise from dialect-specific, systematic effects of the surrounding phonetic environment (e.g., distribution of /æ/-tensing in the mid-Atlantic American English dialect[1,2]). Phonotactic environment can also affect vowel production; southern French dialects have neutralized /e/ and /ɛ/ in word-final context; listeners who speak this dialect have difficulty in perceiving the contrast /e-ɛ/, which is preserved word-finally in Standard French.[3,4] The present study analyzed Canadian French (CF) vowel productions to (1) explore the effect of preceding consonantal context (if any) on vowel production and to (2) quantify the degree to which the vowels /e/ and /ɛ/ were differentiated in word-final context.

1.1 The CF vowel system

CF has eight phonemically contrastive oral vowels in word-final position: /i y ø e ɛ o u a/. Of particular interest was determining whether the /e-ɛ/ distinction is maintained in word-final position. Certain French mid-vowels tend to occur in open syllables (e.g., /e/), whereas others most often occur in closed syllables (e.g., /ɛ/). This phonotactic distribution is referred to as *loi de position*.[5,6] However, both /e/ and /ɛ/ can occur in open syllables word-finally in CF.[7,8] This contrast not only distinguishes lexical minimal pairs [e.g., *fée* /fe/ "fairy" vs *fais* /fɛ/ "(I) make"], but is also used as a morphological suffix (e.g., *je parlerai* /e/ "I will talk" vs *je parlerais* /ɛ/ "I would talk"). Although previous research has shown the maintenance of this contrast when produced in isolated words,[8] it is possible that the /e-ɛ/ contrast is more neutralized when produced in sentential context. The goal of this study was to determine how reliably /e-ɛ/ is differentiated in various phonemic contexts when produced in words embedded in sentences.

2. Methodology: Participants and stimulus materials

Three men and one woman were recorded. All participants were French-dominant, living in Montréal and reporting CF as their first language. Speaker A was born in Jonquiere, CA; the other three were born in Montréal. Speakers A (male, 24 yrs.) and B (female, 18 yrs.) self-identified as monolingual French speakers. Speaker C (male, 32 yrs.) identified as a balanced bilingual. Speaker D (male, 22 yrs.) identified as bilingual with accented English. Participants were compensated for their time.

Target words were embedded in one of three carrier phrases: "*Je dis le mot* _____ *pour lui./tout le temps./que j'ai oublié*." (trans: "I say, the word _____ for

[a)]Author to whom correspondence should be addressed.

him./all the time./that I forgot.") Real words were chosen by finding monosyllabic words with one of the eight vowels in word-final stressed position, preceded by either a labial, coronal, or back consonant. Disyllabic words were used when no monosyllabic word was possible. Nonwords were constructed with the form /ʒɪsCV/, where each vowel was preceded by either a labial, coronal, or back consonant.

In order to explore further the stability of /e-ɛ/ as both a lexical and morphological contrast, four morphological minimal verb pairs were also recorded (three real, one nonword). The nonsense verb /ʒɪste/ was created, based on the form of the most common French verb conjugation. Finally, due to the fact that the vowels in this morphological context are always preceded by the consonant /ʁ/, two monosyllabic words with word-final /e/ or /ɛ/ preceded by /ʁ/ were also included.

3. Procedures: Recording and acoustical analysis

Participants were recorded in a sound-treated room at McGill University, using a Marantz PMD671 solid-state recorder (D&M Holdings Inc., Kanagawa, Japan) and Shure SM81 cardioid condenser microphone with pop filter (Shure Inc., Niles, IL). The built-in 18 dB/octave roll-off high pass filter of the microphone was employed to minimize background noise below 100 Hz. The recordings were digitized at a 44.1 kHz sampling rate, 16-bit resolution. The recording sessions were conducted in French. Two repetitions of each sentence were read in random order, blocked by word type (i.e., real word, followed by nonword).

The participants' productions were downsampled to 22.05 kHz., using *Cool Edit 2000* software.[9] A total of 144 tokens were recorded for each vowel, across speaker (264 tokens for /e/ and /ɛ/). Eight tokens were removed due to obvious mispronunciations of the target vowel. The final vowel was analyzed using LPC analysis in MATLAB, using a 512 pt. window size (approximately 25 ms) and 24 coefficients, with manual correction when necessary. The onset of the final vowel was defined by onset of voicing, determined by vertical striation in the spectrogram and periodic energy in the waveform. The offset was defined by the cessation of upper formant energy. The first three formant measurements at the midpoint of the vocalic nucleus were used for analysis. Although Parisian French listeners use duration minimally as a cue in discriminating vowels,[10,11] listeners who speak other French dialects have been shown to use duration as a phonemic cue.[11] With this in mind, vocalic duration (i.e., offset-offset measures) was also analyzed.

4. Results

4.1 Effect of lexical status and phonetic context

Table 1 lists the mean *F*1-*F*2-*F*3 measures for each vowel and Fig. 1 shows the variation of vowel production by speaker. These patterns are consistent with previous research.[8] Linear discriminant analysis (LDA) is a multivariate regression technique that was used to determine whether there were any differences in vowel classification between real and nonwords.[12] The percentage of correct classification quantifies the discreetness of categories within the dataset, based on the parameters included in the model. In addition, the LDA model may be used to classify a separate set of data, using the parameters estimated from the original set of data. These measures are considered below.

In the following discriminant analyses, the vowel categories were used as the classification labels. The *F*1–*F*2–*F*3 measurements at the 50% point, as well as vocalic duration, were used as the independent variables. In comparing the production of real words to nonwords, the discriminant analysis revealed that these productions were relatively similar. A discriminant analysis modeled on the real words across the three

Table 1. Mean *F*1-*F*2-*F*3 measures (Barks), collapsed across phonetic context, and word type.

	Speaker A			Speaker B (f)			Speaker C			Speaker D		
	*F*1	*F*2	*F*3	*F*1	*F*2	*F*3	*F*1	*F*2	*F*3	*F*1	*F*2	*F*3
/i/	2.5	12.8	14.8	3.0	14.0	15.8	1.8	13.8	15.3	2.5	13.0	16.1
/e/	3.2	12.7	14.5	3.3	14.1	15.4	2.4	13.8	14.9	3.2	13.6	15.0
/y/	2.6	11.9	13.3	3.0	12.9	14.1	2.0	12.3	13.3	2.7	12.2	13.3
/ø/	3.3	11.2	13.3	3.9	11.5	13.3	2.7	10.8	12.9	3.3	10.8	12.9
/u/	2.6	7.1	13.7	3.0	8.3	14.3	2.2	6.0	13.7	2.4	6.7	13.9
/o/	3.3	7.1	13.9	4.0	7.6	14.6	2.7	5.7	14.2	3.1	5.7	15.0

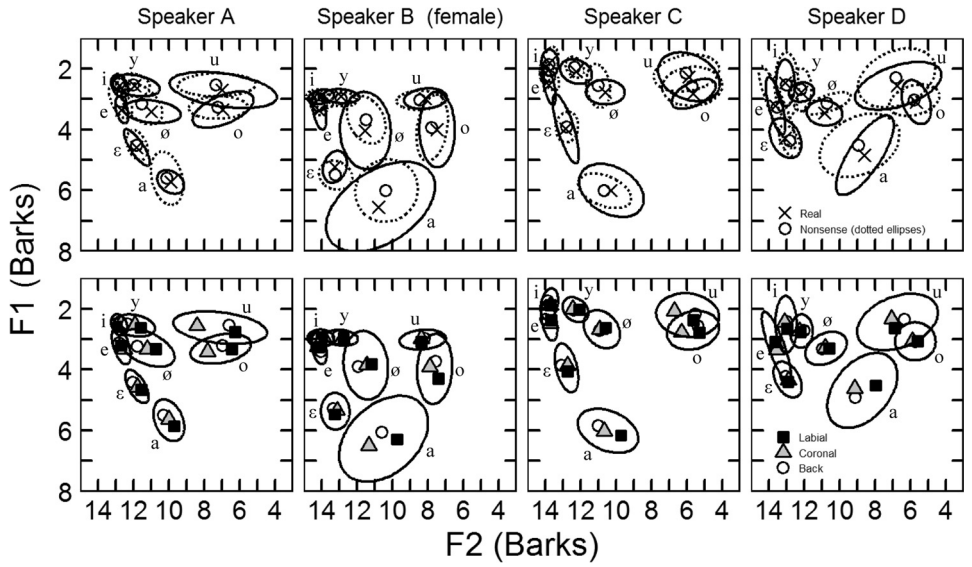

Fig. 1. Comparison of mean formant values *F*1-*F*2 vowel space by preceding consonantal context (top row) and real vs nonwords for each speaker (bottom row). Ellipses = 1 standard deviation (SD) of vowel productions.

male speakers was able to correctly classify 92% of these tokens. The LDA models of real word tokens for each speaker were as follows: for Speaker A, 97% real-word model classified 90% of the nonword stimuli, Speaker B: 94% the model classified 94% of the nonword stimuli, Speaker C: 95% the model classified 91% of the nonword stimuli, Speaker D: 97% the model classified 96% of the nonword stimuli. Similar classification accuracies were observed when real words were classified based on nonword models.

Regarding duration differences in real word vs nonword productions, paired-samples *t*-tests were calculated for each speaker to analyze vowel duration. The durations were paired such that the mean duration across the two tokens of the real word tokens of a vowel were matched to the mean duration of the nonsense tokens of the same preceding context. All speakers produced longer vowels for real words than nonwords, controlling for context and vowel, although this pattern did not reach significance for Speaker B [Speaker A: $M = 6$ ms, $t(23) = 3.70$, $p = 0.001$; Speaker B: $M = 5$ ms, $t(23) = 1.94$, $p = 0.064$; Speaker C: $M = 11$ ms, $t(22) = 5.55$, $p < 0.001$; Speaker D: $M = 10$ ms, $t(23) = 3.93$, $p = 0.001$]. This finding was consistent with previous research, which found that word-final vowels in monosyllabic words are longer than final vowels in disyllabic words.[8,13] More important, the speakers did not slow down when producing the novel nonwords, providing further evidence that production of nonwords was relatively natural and not overemphasized.

Regarding the effect of phonetic context, there were overall only minimal effects of preceding context on vocalic midpoint formant values, mainly limited to relative fronting of back vowels in coronal context and backing of /a/ in labial context. For the most part, the gestural target was reached by vocalic midpoint. Spectral distinctions between /e-ɛ/ were maintained for all speakers in all contexts.

4.2 Spectral clustering of high front-unrounded /i e/, front-rounded /y ø/, and back-rounded /u o/ vowels

Although previous studies have suggested that French /u/ and /y/ are spectrally similar when exploring the vowel space in a traditional *F*1–*F*2 representation,[14] the vowels for the speakers analyzed here clustered in the following way in *F*1–*F*2–*F*3 space: front-unrounded vowels had relatively high *F*2 and high *F*3 values, back-rounded vowels had a relatively lower *F*2, but high *F*3 values, and front-rounded vowels had *F*3 values lower than both back-rounded and front-unrounded vowels, with *F*2 values between the back-rounded and front-unrounded vowels (see Table 1), suggesting that in the productions reported here, CF front-rounded vowels are not spectrally similar to front-unrounded vowels.

4.3 Production of French /e-ɛ/ in lexical and morphological contrasts

Figure 2 illustrates the degree of /e-ɛ/ spectral differentiation by target word type for each speaker. All subjects reliably maintained the /e-ɛ/ distinction spectrally in lexical

tokens when preceded by a labial, coronal or back stop. Speaker D did show some overlap in production, due to a large variability in *F*1 values for /e/ productions (gray ellipses, Fig. 2). A Wilcoxon Signed-Ranks test (nonparametric equivalent of a paired *t*-test) was used to compare the durations of /e/ and /ɛ/ in lexical tokens for each speaker. The mean duration of /e-ɛ/ productions in real and nonsense words were paired by preceding context (labial, coronal, or back stop). The results of a Wilcoxon Signed-Ranks test revealed that all speakers produced /ɛ/ with a significantly longer duration of approximately 15 ms, compared to /e/ [Speakers A, C, D: $N=6$, $T+=6$, $z=-2.20$, $p=0.028$, Speaker B: $N=5$, $T+=5$, $z=-2.02$, $p=0.043$]. A Kruskal-Wallis test (nonparametric equivalent of a one-way analysis of variance) including all speakers was calculated on the duration differences between /e/ and /ɛ/ to determine whether the amount of difference varied depending on the preceding stop. The result was not significant; the duration of the /e-ɛ/ did not vary as a function of the preceding consonantal context [$\chi^2=4.43$, $df=2$, $p=0.109$].

In order to compare differences in production of /e-ɛ/ in lexical vs morphological words, it was first necessary to explore differences in production of this contrast in lexical tokens when preceded by /ʁ/. This aided in differentiating production patterns that were related to word type from patterns that were related to phonetic environment. For the lexical tokens, both monolingual Speakers A and B exhibited little spectral overlap in /e-ɛ/ productions preceded by /ʁ/. Speaker A showed very little coarticulation; however, Speaker B exhibited spectral lowering of /e/ and backing of /ɛ/. The two bilingual speakers exhibited more spectral overlap, but in different patterns: Speaker C had more variability in /ɛ/ productions, whereas Speaker D produced /e/ with more variability (Fig. 2, dashed ellipses). Although all speakers produced a reliable duration difference when preceded by a stop, this was not the case when preceded by /ʁ/. The Kruskal-Wallis test of differences in duration as a function of preceding context was recalculated, this time including lexical tokens preceded by /ʁ/. The results were significant [$\chi^2=13.09$, $df=3$, $p=0.004$]; duration differences in productions of /e-ɛ/ preceded by /ʁ/ were ranked as the smallest for all four speakers. See Table 2 for mean durations of /e/ and /ɛ/ by preceding context.

For production of /e-ɛ/ in morphological minimal pairs, the two monolingual speakers maintained a reliable spectral contrast. The two bilinguals again showed a different pattern; Speaker C exhibited some overlap, again mainly due to variability in production of /ɛ/. Speaker D completely neutralized the contrast for morphological minimal pairs; /e/ and /ɛ/ were neutralized to /ɛ/ (see Fig. 2, solid, unfilled ellipses). The durations of /e-ɛ/ in morphological tokens were very similar to the durations in lexical tokens when preceded by /ʁ/, again exhibiting minimal duration differences between /e/ and /ɛ/.

To summarize, in lexical minimal pairs, the /e-ɛ/ contrast was differentiated both spectrally and temporally by all speakers, in all preceding phonetic contexts. /e/ was spectrally lower and longer in duration when preceded by /ʁ/ than in other contexts. Three of the four participants maintained a stable distinction in morphosyntactic contexts. Monolingual subjects exhibited the best retention of the distinction, whereas the bilingual participants had partially or completely overlapping distributions, unexpectedly merging /e-ɛ/ to /ɛ/. This pattern is different from that observed in Southern French, where /ɛ/ has merged to /e/.[3,4]

5. Discussion

The eight vowels of CF were well distinguished by all speakers in word-final position. Specifically, the /e-ɛ/ contrast was maintained by all speakers in lexical minimal pairs,

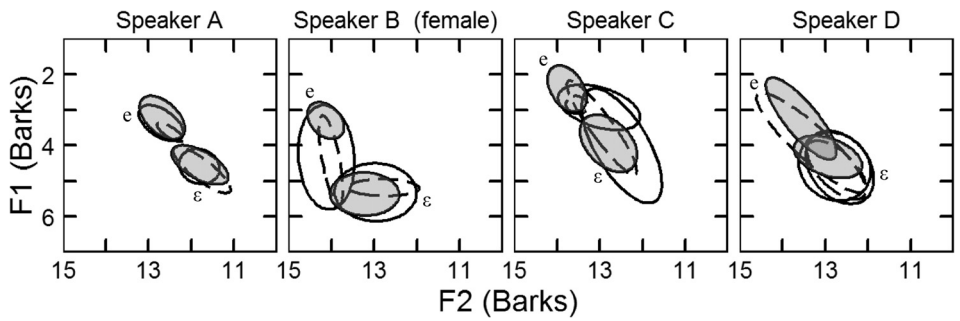

Fig. 2. Comparison of /e-ɛ/ variability in *F*1-*F*2 vowel space for each speaker. Filled ellipse = lexical. Dashed = /ʁ/ lexical. Solid unfilled = morphological. Ellipses = 1 SD of vowel productions.

Table 2. Mean durations (in ms) for /e/ and /ɛ/ productions, by preceding phonetic context. First four columns: lexical tokens, fifth column: morphological tokens (vowel preceded by /ʁ/).

Speaker	Labial	Coronal	/e/ Back	/ʁ/	Verb	Labial	Coronal	/ɛ/ Back	/ʁ/	Verb
A	88	78	77	93	94	98	93	99	94	96
B (f)	92	82	93	108	105	95	101	115	110	114
C	102	101	106	114	116	120	116	117	114	115
D	106	99	101	114	110	117	109	126	116	115

differing both spectrally and by duration. The monolingual speakers also reliably maintained this contrast in morphosyntactic minimal pairs. The two bilingual speakers had either partial or complete overlap, suggesting perhaps that this contrast is not as stable for the bilingual speakers. Regarding possible temporal/spectral differences between real and nonwords, there were no major differences between real and nonwords.

The results of this study contribute to predictions made in the body of literature investigating cross-language speech perception of vowels, particularly with respect to difficulty in discriminating /e-ɛ/. Previous studies investigating /e-ɛ/ in Catalan have demonstrated that this contrast can be difficult to discriminate by Spanish-dominant bilinguals[15,16] and that perception of this contrast by Spanish-Catalan bilingual infants diverges from monolingual Catalan infants as early as eight months.[17] French vowels /e/ and /ɛ/ are difficult to discriminate by listeners of French dialects that have neutralized this context[3,4] /e-ɛ/ are sometimes confused for each other by English listeners, especially in open-syllable context,[5] and French /ɛ/ is sometimes perceived as most similar to English /e/.[18] Despite the fact that /e/ and /ɛ/ are phonemic in both English and French, French /e/ and /ɛ/ tend to be spectrally higher than in English[14] and /ɛ/ can occur word-finally in French, but not in English. Finally, the speakers in the present study produced /ɛ/ with a duration that was longer than /e/ when preceded by a stop, or roughly equivalent when preceded by /ʁ/. This pattern contrasts with English, where /ɛ/ is a lax vowel and therefore shorter in duration than /e/, particularly in stressed syllables.[19,20] English listeners have been demonstrated to use English expectations of duration in perceiving French vowels, even with having experience in speaking French,[10] and could potentially have difficulty in discriminating CF /e-ɛ/ in word-final position; this contrast will be of particular interest in future cross-language perceptual studies involving CF second-language learners.

Acknowledgments

The authors thank Linda Polka and McGill University for providing laboratory space and resources. Portions of this work were presented at Acoustical Society of America meetings in June 2008, and May 2009. Supported by NIH Grant Nos. F31 DC008075 and 2 T32 HD049899-06.

References and links

[1]S. Ash, "The distribution of a phonemic split in the Mid-Atlantic region: Yet more on short a," U. Penn Working Papers Ling. **8**(3), 1–15 (2002).

[2]W. Labov, "Exact description of the speech community: Short a in Philadelphia," in *Language Change and Variation*, edited by R. W. Fasold and D. Schiffrin (Georgetown University Press, Washington, DC, 1989), pp. 1–57.

[3]S. Dufour, N. Nguyen, and U. H. Frauenfelder, "The perception of phonemic contrasts in a non-native dialect," J. Acoust. Soc. Am. **121**(4), EL131–EL136 (2007).

[4]S. Dufour, N. Nguyen, and U. H. Frauenfelder, "Does training on a phonemic contrast absent in the listener's dialect influence word recognition?," J. Acoust. Soc. Am. **128**(1), EL43–EL48 (2010).

[5]T. L. Gottfried, "Effects of consonant context on the perception of French vowels," J. Phonetics **12**(2), 91–114 (1984).

[6]D. C. Walker, *The Pronunciation of Canadian French* (Ottawa University Press, Ottawa, 1984), pp. 32–89.

[7]J. Boshoer and R. Mougeon, "Les réalisations par des élèves d'immersion de [e] et [ɛ] en position finale accentuée dans les formes verbales" ("The achievements by immersion students [e] and [ɛ] in stressed final position in the verbal forms"), in *French Studies* (York University, North York, Ontario, 2004).

[8]P. Martin, "Le système vocalique du français du Québec. De l'acoustique à la phonologie" ("The vowel system of French Quebec. From the acoustic phonology"), La Linguistique **38**(2), 71–88 (2002).

[9]D. Johnston, Cool Edit Pro, Syntrillium Software Corporation, Phoenix, Arizona (1999).

[10]T. L. Gottfried and P. S. Beddor, "Perception of temporal and spectral information in French vowels," Lang. Speech **31**, pt. 1(3), 51–75 (1988).

[11]J. L. Miller, M. Mondini, F. Grosjean, and J.-Y. Dommergues, "Dialect effects in speech perception: The role of vowel duration in Parisian French and Swiss French," Lang. Speech **54**(4), 467–485 (2011).
[12]W. R. Klecka, *Discriminant Analysis* (Sage Publication, Newbury Park, CA, 1980).
[13]B. Lindblom, "Temporal organization of syllable production," STL-QPSR **9**(2–3), 1–5 (1968). Retrieved from http://www.speech.kth.se/prod/publications/files/qpsr/1968/1968_9_2-3_001-005.pdf (Last viewed June 15, 2015).
[14]W. Strange, A. Weber, E. S. Levy, V. Shafiro, M. Hisagi, and K. Nishi, "Acoustic variability within and across German, French, and American English vowels: Phonetic context effects," J. Acoust. Soc. Am. **122**(2), 1111–1129 (2007).
[15]C. Pallier, L. Bosch, and N. Sebastián-Gallés, "A limit on behavioral plasticity in speech perception," Cognition **64**(3), B9–B17 (1997).
[16]C. Pallier, A. Colomé, and N. Sebastián-Gallés, "The influence of native-language phonology on lexical access: Exemplar-based vs. abstract lexical entries," Psychol. Sci. **12**(6), 445–449 (2001).
[17]L. Bosch and N. Sebastián-Gallés, "Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life," Lang. Speech **46**(2–3), 217–243 (2003).
[18]W. Strange, E. S. Levy, and F. Law II, "Cross-language categorization of French and German vowels by naive American listeners," J. Acoust. Soc. Am. **126**(3), 1461–1476 (2009).
[19]E. Jacewicz, R. A. Fox, and J. Salmons, "Vowel duration in three American English dialects," Am. Speech **82**(4), 367–385 (2007).
[20]M. Picard, *Phonetics and Phonology for ESL and TESL Teachers* (Concordia University, Montreal, 2001).